



Il pontefice, lo scienziato e il prete della Silicon Valley Alle origini di un'alleanza

Olah, uno dei padri di Claude, era alla presentazione in Vaticano
Il legame con Prevoist nato grazie a un informatico diventato sacerdote

Nella sua prima enciclica, *Magnifica Humanitas*, Leone XIV riconosce che l'ascesa dell'intelligenza artificiale rischia di concentrare ricchezza e potere «nelle mani di pochi». Il riferimento è alle grandi società tecnologiche della Silicon Valley, capaci di sviluppare sistemi diffusi in tutto il mondo e di accumulare valore in una geografia ristretta, rappresentata simbolicamente da San Francisco. È qui che ha sede Anthropic, società fondata dai fratelli italoamericani Daniela e Dario Amodei e diventata in breve tempo un colosso tech del valore di oltre 350 miliardi di dollari grazie a Claude, la sua IA capace di svolgere lavori in autonomia e stabilire obiettivi militari. Anthropic, insomma, appartiene a quel ristretto gruppo di aziende che l'enciclica invita a guardare con maggiore attenzione: soggetti privati capaci di orientare una tecnologia destinata a incidere sulla vita di miliardi di persone.

È quantomeno singolare, dunque, che accanto ai cardinali e ai teologi chiamati a commentare il testo papale, in Vaticano sia comparso il 34enne canadese Christopher Olah, cofondatore di Anthropic e capo della ricerca sull'interpretabilità. Olah non è un manager simbolo della Silicon Valley. Il suo lavoro consiste nel tentativo di aprire la "scatola nera" delle reti neurali, cioè di capire come un sistema di intelligenza artificiale arrivi a produrre una risposta. Tra Leone XIV e Olah ha fatto da ponte padre Brendan McGuire, sacerdote irlandese di sessant'anni e parroco di St. Simon a Los Altos, nel cuore della Silicon Valley. McGuire lavorava nell'industria tecnologica prima di entrare nella Chiesa. Ordinato sacerdote, non ha mai reciso il legame con la Valley. È diventato anzi uno dei punti di riferimento religiosi di ingegneri e ricercatori che, a un certo punto, hanno cominciato a porsi domande morali mentre cercavano di costruire un cervello artificiale.

LA STORIA
di PIER LUIGI PISA

A marzo scorso, secondo il *Washington Post*, McGuire era tra i leader cristiani che Anthropic ha ospitato nei suoi uffici. Nel corso dell'incontro, organizzato in gran parte proprio da Olah, gli ingegneri dell'azienda avrebbero chiesto indicazioni sullo sviluppo morale e spirituale di Claude, compresa la possibilità che il chatbot potesse essere considerato



Christopher Olah, ricercatore e cofondatore di Anthropic, società che produce intelligenza artificiale; in alto, Olah con papa Leone XIV Sotto, padre Brendan McGuire

"Abbiamo bisogno di voci morali che gli incentivi non possano piegare"

un "figlio di Dio". È stato quello il primo passo di Olah verso la Santa Sede. Ieri, in Vaticano, si è trovato davanti un Papa che affronta l'intelligenza artificiale come la questione sociale del XXI secolo. Per il pontefice, il problema non riguarda solo l'uso buono o cattivo di una tecnologia, ma anche il modo in cui questa viene progettata e le "scelte e priorità" che incorpora. Su questo aspetto si concentra proprio Olah: rendere più comprensibili i sistemi; capire perché producono certi risultati e rendere possibile una responsabilità effettiva. Sul piano teorico, le motivazioni di Leone XIV e Olah sono molto simili. Ma è lo stesso cofondatore di Anthropic ad ammettere, nel suo intervento al Vaticano, che per un'azienda di intelligenza artificiale «fare la cosa giusta» può essere complesso. In gioco ci sono «incentivi», «vincoli» e almeno due pressioni decisive: «quella geopolitica», con un'America decisa a non farsi superare dalla Cina nella corsa all'IA, e quella economica, legata alla necessità di restare «commercialmente sostenibili». Olah ha concluso il suo discorso affermando: «Abbiamo bisogno di voci morali che gli incentivi non possano piegare». Viene da chiedersi se, dopo l'incontro in Vaticano, quella voce morale finisca per rafforzare proprio uno degli attori che dovrebbe contribuire a sorvegliare. Olah torna infatti negli Stati Uniti con una credibilità accresciuta dall'effetto Santa Sede.

E con lui Anthropic, che può presentare la sicurezza dell'IA come vocazione prima ancora che come strategia aziendale. Per un'impresa ormai esposta ai contratti militari, alla competizione feroce della Silicon Valley e a una crescente frustrazione delle persone che temono di essere sostituite dalle macchine, il richiamo del Papa alla custodia della persona umana offre una legittimazione che nessuna campagna di comunicazione potrebbe costruire da sola.

IL COMMENTO
di ANTONIO SPADARO

La risposta della dottrina per migliorare l'algoritmo

La mossa decisiva della *Magnifica Humanitas* – la prima enciclica di Leone XIV, firmata il 15 maggio 2026 – non è dove il lettore se la aspetta. Non è nella denuncia dei monopoli tecnologici, né nella richiesta di regolazione degli algoritmi, né nella condanna dei sistemi d'arma autonomi: tutte cose che il documento fa, e fa molto bene. È in una domanda che nessun *white paper* della Silicon Valley e nessun regolamento europeo avrebbero mai formulato: che cosa amiamo davvero? Il Papa cita Agostino – «due amori fecero due città» – e sposta l'intero dibattito sull'intelligenza artificiale sul terreno del desiderio, della direzione interiore delle scelte, di ciò che una civiltà considera degno di essere perseguito prima ancora di decidere come perseguirlo.

La pubblicazione avviene a centotrentacinque anni dalla *Rerum novarum* di Leone XIII, l'enciclica che nel 1891 mise la questione operaia al centro della riflessione della Chiesa. Oggi il nuovo Leone reclama lo stesso ruolo per la questione algoritmica. Come il primo intervenne sulla dignità del lavoro di fronte allo sconvolgimento della fabbrica, così Prevoist interviene sulla dignità della persona di fronte allo sconvolgimento dei dati. Ma il salto vero è metodologico: come Sconosciuto con la *Laudato si'* aveva proiettato la questione ecologica dentro la Dottrina sociale della Chiesa, allo stesso modo Leone agisce per l'intelligenza artificiale, ponendola come nuovo centro di gravità – una trasformazione che interpella dall'interno le categorie della dottrina sociale e ne esige un ripensamento. Non un'appendice: il capitolo che riscrive tutti gli altri.

L'architettura del testo poggia su due figure bibliche che funzionano come modelli politici. Babele: lingua unica, sapere totalizzante, potere centralizzato che si edifica sull'uniformità. Leone ne fa la diagnosi di quella che chiama «sindrome di Babele» – la pretesa di tradurre tutto, compreso il mistero irriducibile della persona, in dati e prestazioni. All'opposto, Neemia: l'ebreo in esilio che torna nella Gerusalemme distrutta non impone un piano dall'alto, osserva le rovine, convoca le famiglie, affida a ciascuna un tratto di muro. La città rinasce per responsabilità condivisa. La scelta non è tra tecnologia sì e tecnologia no, ma tra due logiche di potere: concentrarlo o distribuirlo, omologare o comporre.

Le innovazioni dottrinali sono concrete. La destinazione universale dei beni – applicata dalla tradizione cattolica alla terra – viene estesa ai brevetti, agli algoritmi, ai dati. Il Papa applica all'economia degli algoritmi la stessa logica che la Chiesa applica da secoli al latifondo: quando quei beni restano in poche mani, la disuguaglianza si struttura. La sussidiarietà, che proteggeva i corpi intermedi dallo Stato, viene ribaltata: nell'ecosistema digitale chi assorbe competenze non è il governo ma sono le piattaforme, che definiscono accesso, visibilità, relazioni. Leone nomina i «nuovi monopoli dell'IA» e un'asimmetria epistemica, economica e politica incompatibile con il bene comune. Il verbo che sceglie – «disarmare» l'intelligenza artificiale – sposta il discorso dalla *compliance* alla struttura del potere globale.

Certo, la metafora di Babele presuppone un progetto unitario, mentre la rivoluzione digitale è policentrica; la destinazione universale dei dati è un principio potente, e la sua traduzione in governance resta da costruire. Il punto in cui l'enciclica diventa particolarmente incisiva è quando legge l'intelligenza artificiale come questione che riguarda la struttura stessa dell'intelligenza umana – la sua capacità di conoscere, di desiderare, di sbagliare, di ricominciare. Il paradigma tecnocratico non è solo un errore di governance: è una deformazione dell'immaginario, la sostituzione silenziosa della sapienza con l'efficienza, della relazione con la prestazione, del giudizio con il calcolo.

Da qui le pagine sulla simulazione della relazione: quando la parola viene simulata non costruisce un legame, ma una «parvenza» – e il danno non è che qualcuno scambi la macchina per una persona, ma che perda il desiderio stesso di cercare l'altro. Sul limite: ciò che appare come fragilità viene oggi letto come difetto da correggere, e l'enciclica obietta che l'umano non fiorisce malgrado il limite ma attraverso il limite – posizione che sfida frontalmente il transumanesimo. Sull'errore: per un algoritmo è qualcosa da eliminare; per un essere umano può essere l'inizio di una trasformazione radicale. Il futuro di una persona non è calcolabile. Viene in mente il verso dichiarativo di Emily Dickinson: Io abito la possibilità.

L'enciclica propone spazi deliberati di vuoto in cui il pensiero maturi senza la stampella del calcolo. Non è luddismo. È la constatazione che un'attenzione permanentemente mediata dall'algoritmo finisce per atrofizzarsi, e che proteggere la capacità di pensare in proprio è diventata una questione politica.

La *Magnifica Humanitas* si chiude con il *Magnificat* di Maria e con Tolkien: «Non tocca a noi dominare tutte le maree del mondo, ma fare il possibile per la salvezza degli anni in cui viviamo». Che un'enciclica sull'intelligenza artificiale trovi la sua parola definitiva in un inno del I secolo e in un romanzo del Novecento dice molto sulla visione di Leone XIV. La risposta alla sfida tecnologica non è un algoritmo migliore. È una diversa qualità dello sguardo sulla qualità umana del progresso, che racconto e poesia custodiscono infallibilmente. E il Papa agostinianamente chiede al suo lettore: «Che cosa ami davvero?».